The influence of feature-based information in the age processing of unfamiliar faces

Article (Published Version)

Perception, 1998, volume 27, pages 295–312

The influence of feature-based information in the age processing of unfamiliar faces

Patricia A George, Graham J Hole
School of Cognitive and Computing Sciences, University of Sussex, Falmer, Brighton BN1 9QH, Sussex, UK; e-mail: pgeorge@cogs.susx.ac.uk; grahamh@cogs.susx.ac.uk
Received 4 June 1996, in revised form 20 January 1998

Abstract. The influence of the internal features (eyes, nose, and mouth) in the age processing of unfamiliar faces was examined. Younger and older versions of the faces of six individuals (covering three different age ranges, from infancy to maturity) were used as donor stimuli. For each individual in turn, the effects on age estimates of placing older features in the younger face version (or vice versa) were investigated. Age estimates were heavily influenced by the age of the internal facial features. Experiment 2 replicated these effects with a larger number of faces within a narrower age range (after growth is complete and before major skin changes have occurred). Taken together, these two experiments show that the internal facial features may be influential in conveying age information to the perceiver. However, the mechanisms by which features exert their influence remain difficult to determine: although age estimates might be based on local information from the features themselves, an alternative possibility is that featural changes indirectly influence age estimates by altering the global three-dimensional shape of the head.

1 Introduction

Faces convey a richness of perceptual information: they facilitate communication and convey information about a person's emotional state, identity, gender, and age. A study by Shepherd et al (1981) demonstrates the importance of age as a facial characteristic: when subjects were asked to categorise faces, three dimensions (hair, face shape, and apparent age) seemed to account for all perceived variation between faces.

The dimensions of age and face shape (discussed below) highlight perceived age as a potentially important attribute in the initial encoding of faces. The work of dentists and plastic surgeons (eg Enlow 1968, 1982; Viidik 1973) is concerned with the way the individual face physically changes over time. They suggest that the ageing face changes in differing ways depending on the age of the face itself; ageing not only affects the growth of the cranium as it changes in size and shape, but also produces more subtle changes in skin texture and muscle tone.

There have been a number of studies that have explored perception of the ageing face from the perspective of growth. Pittenger, Shaw, Mark, and their colleagues (eg Pittenger and Shaw 1975; Pittenger et al 1979; Mark and Todd 1985; see also Bruce et al 1989) have shown that changes in the shape of the skull during growth can be modelled by a mathematical transformation called cardioidal strain. These researchers suggest that sensitivity to the level of cardioidal strain underlies judgments of facial age. (In this paper, we will use the term 'ageing' to refer to changes occurring throughout life, ie from birth to death. 'Ageing' therefore subsumes the period of 'growth', a term we shall use to refer to the period from birth to maturity.)

The type of global changes in the human head suggested by Enlow (1982), and the growth transformation developed by Pittenger, Shaw, and their colleagues, emphasise faces as being three-dimensional shapes (eg Mark and Todd 1983). However, the changing shape of the head (as represented, for example, by cardioidal-strain level) cannot be a complete explanation of how people perceive age, since these kinds of changes are only available as a source of age information while the cranium is growing. Since age-discrimination judgments can be made between different-aged adult faces [ie after growth is complete (eg George and Hole 1995)], other sources of information must be available as cues to age.

We previously (George and Hole 1995) explored factors that influenced judgments of age, using—as in the present experiments—stimuli which were more naturalistic than in previous studies (and hence potentially capable of providing a greater variety of cues to age). We found that subjects were highly accurate in judging the age of faces that ranged between 5 and 70 years. To try to establish what cues were important in this task a series of manipulated versions were produced that either reduced or enhanced cues that were thought to carry age information. Comparing the performance between experimental conditions suggests that a rather subtle use of cues is implicated when making judgments of apparent age.

Subjects were influenced by 'pseudo' cardioidal-strain-level cues, in which faces that had been manipulated in a way to make them appear younger (by altering the location of the eyes, nose, and mouth) were perceived as such. However, age-estimation performance was poor for faces that had had all surface detail (eg skin-texture and feature cues) removed but retained information about cardioidal-strain level and configurational properties. This decrement in performance would not be expected if cardioidal-strain-level cues were the primary source of age information. Support for the importance of other sources of age information was provided by the fact that age-estimation performance was as accurate as that of controls when only the internal face features were presented. The necessary and sufficient sources of age information that are available in the internal part face could be the details of the features themselves, the spatial arrangements of the features, the available surface information, or any combination of these.

In the present study we investigate whether specific facial features have a role to play in conveying age information. However, this highlights perhaps one of the most difficult and frequently encountered distinctions in research on face processing and recognition that is made concerning the kinds of information that are available in faces—the distinction between feature-based and configurational information. [Various authors have made this distinction in subtly different ways. For example, Carey and Diamond (1977) distinguish between 'component' and 'configural' processing, whereas Rhodes et al (1993) contrast 'isolated-feature' and 'second-order-relational' modes of processing. As a starting point, we will use Bruce's (1988) definition of a feature as "a discrete component part such as a nose" and her definition of configural information as referring to the "spatial interrelationships of facial features" (page 38), while acknowledging that in practice the two sources of information may be interdependent.]

There are at least two problems in testing the relative importance of feature-based versus configural information in face processing. The first is the issue of deciding, in any principled way, what constitutes a 'feature' and what constitutes 'configurational' information. It is routine, both in everyday life and in psychological research, to refer to individual facial features as discrete entities—ie to treat the eyes, nose, and mouth as 'parts' of the face. However, merely because these parts have subjective salience, it does not mean that they are necessarily salient to the visual system in the same way. Nevertheless, there are independent sources of evidence to suggest that facial parts do have some special status in face processing. For example, psychological studies of feature salience (eg Davies, Ellis, and Shepherd 1977, cited in Davies et al 1981; Fraser and Parker 1986) have shown that eyes, noses, and mouths are preferentially attended to by subjects. Eye-movement-recording studies (eg Luria and Strauss 1978) have shown that in tasks involving individual recognition, subjects tend to fixate the central facial area, specifically paying attention to the eyes, nose, and mouth. There is some evidence that, after inversion of a face, configurational changes are less noticeable than changes to isolated features (Bartlett and Searcy 1993; Rhodes et al 1993—but see the discussion in Rhodes 1995). Neurophysiological studies (eg Perrett et al 1987) have demonstrated the existence of cortical cells in the monkey which respond preferentially to individual features (especially the eyes) as well as cells which respond primarily to more 'holistic' (configurational) aspects of faces.

Thus, at least as far as recognition is concerned (that is, allowing for the possibility that other types of face processing, such as those related to age perception or gender categorisation might differ in this respect), there are reasonably good grounds for making a distinction between those parts of a face conventionally thought of as 'features' and other facial aspects such as 'configurational' properties. We are not claiming that other parts of a face, such as the forehead, chin, or cheek region, should not be considered as 'features'. We are using the traditionally defined 'features' as a starting point for two reasons. First, as just mentioned, these seem (on both subjective and empirical grounds) to be influential in face processing. Second, whatever the merits of more or less arbitrarily segmenting a face into its constituent parts, because in previous research the role of these features has been examined it is interesting to know how their influence in age perception compares with their influence in other aspects of face processing (such as recognition of individuals).

The second problem is a more practical one, related to the difficulty of selectively manipulating one of these sources of information without affecting the others. Varying the configurational properties of a face independently of the specific feature information is arguably less problematic than trying to do the opposite. While changing the location of the features does not significantly alter the features themselves, changing specific features affects the configurational properties simply because the features are in a spatial interrelationship with each other. In age-perception research the importance of the spatial interrelationship of facial features has been explored by altering the location of specific features (eg Bruce et al 1991; George and Hole 1995), but the role that the features themselves may have on judgments of apparent age has not been considered. The only study that we are aware of is a study of age perception in 4-year-old children by Jones and Smith (1984). They used a masking technique (selective obliteration of various parts of the face) to see which features were most important for age perception, and found that masking the eyes (but not the hair, mouth, or chin) significantly reduced the accuracy of age judgments. Although this research highlights the importance of the internal features for age perception in children, it remains to be determined precisely what aspect of the eye region is important and whether these effects also occur in adults.

The central question asked in the present study is: do the facial features themselves have an influence on the processing of a face's age? The reason for focusing on facial features in the context of age perception is that they are perceptually salient aspects of the adult face which alter over time in response to biological changes produced by ageing. Eyes get smaller and sink deeper into their orbits; hair thins and greys; ears and noses lengthen [reflecting the continuous growth of cartilage throughout life (Smith 1978)]; eyelids and the corners of the mouth drop and sag owing to changes in skin elasticity (see the review in Liggett 1974; see also Enlow 1982). Many of these changes will of course also produce changes which might be regarded as configurational in nature; for example, smaller eyes and larger noses will lead to changes in the spatial relationship between these features, and thinning hair will lead to a higher hairline and hence a larger forehead (although, as one reviewer pointed out, the 'forehead' might be considered a 'feature' in its own right. Such arguments highlight the ambiguity of making too hard and fast a dichotomy between 'features' and 'configurations').

One way to investigate the influence of feature-based information on age perception is to compare age-estimation performance between control faces and faces that have substituted eyes, nose, and mouth of a different age in the same location as the control faces. However, there may be problems in substituting features. As Enlow (1982) has pointed out, certain head shapes and features tend to co-occur, to the extent that one can identify characteristic facial 'types' such as the leptoprosopic (long and thin), euryprosopic (short and fat), and those of gender. When choosing features for substitution, ideally the substituted features should be appropriate, in Enlow's terms, for that particular type of face.

In the present study we ensured that the features used for substitution were appropriate, by using the same face at two different ages: younger or older features were transposed onto an older or younger version of the same individual's face, in the same location as the original features. Two different-age full-face photographs for each of six individuals were used as the starting point for the stimuli in experiment 1. For each individual four versions were produced, two unmanipulated versions 'younger' and 'older' (original younger and older face) and two manipulated versions: younger face with older features (yfof) and older face with younger features (ofyf). The subjects' task was to estimate the apparent age of each of the resultant twenty-four faces. By comparing age-estimation performance between experimental conditions, a measure of the influence of the features themselves can be made.

Our predictions are as follows. If feature information is influential in the processing of a face's age, then a younger face with older features should appear older than the corresponding younger (unaltered) face from which it was derived. Conversely, an older face with younger features should appear younger than the corresponding older (unaltered) face. If age judgments depend solely on factors other than the internal features (for example, head shape, configurational properties, and/or surface information), feature manipulations should have little effect: age estimates should be based on the recipient face rather than on the details of the features transplanted into it.

2 Experiment 1

2.1 *Method*

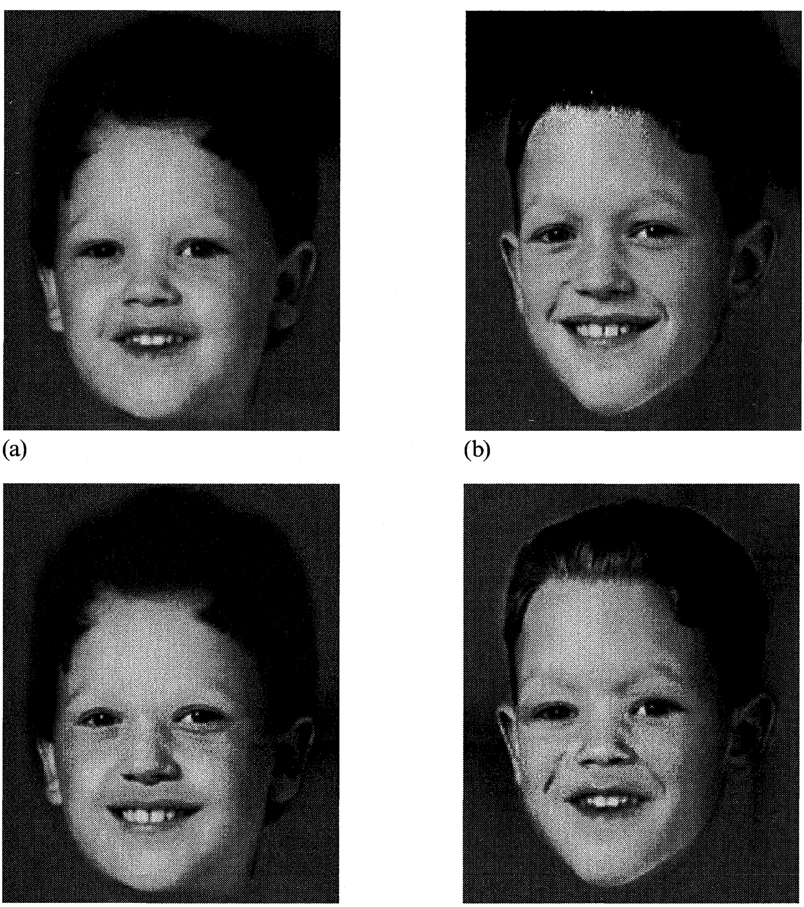
2.1.1 *Subjects.* Forty subjects aged between 18 and 40 years took part. Half of the subjects were male, half were female. All subjects were unaware of the purpose of the experiment.

2.1.2 *Design.* A within-subjects design was used, with all subjects participating in all four conditions of the experiment. There were two independent variables. The first was the face identity. There were three age ranges, with a male and female face in each: infant to young child (henceforth referred to as 'male 1–5 years' and 'female 1–5 years'), young child to older child ('male 4–10 years' and 'female 3–9 years'), and older child/young adult to older adult ('male 10–40 years' and 'female 20–65 years'). The second independent variable was the version of the target face. Four different versions of each face were used, two derived from photographs (one recent and one of the person at a younger age) and two consisting of manipulated versions of these photographs. The four versions were: (a) younger (henceforth referred to as 'younger'); (b) younger face with older features (yfof); (c) older face with younger features (ofyf); and (d) older ('older'). (See below for definitions.) The dependent variable was the subject's estimate of the age of each face, in years.

2.1.3 *Apparatus and materials.* Two full-face photographs (one younger and one older) were obtained for each of six individuals. These were scanned into Adobe Photoshop image-editing software on a Power Macintosh 7100/66 by using a Macintosh OneScanner set for 256 grey levels and 150 dots inch^{-1}. For each of the individuals a set of four stimulus images was produced. Each set consisted of the original younger and older faces plus two variants, yfof and ofyf. After manipulation the images were output as A4-sized prints on a BLP Eclipse 8 laser printer. An example of a set of four versions can be seen in figure 1. The different versions were as follows:

(a) The 'younger' version was an unaltered print from the original (younger) scanned photograph.
(b) The 'older' version was an unaltered print from the original (older) scanned photograph.
(c) The yfof (young face with older features) version was intended to present all the information available in the younger face except for specific features. The resultant facial image placed the selected features in opposition to all other sources of information. The eyes, nose, and mouth from the older version were pasted onto the younger version, in the same location as the features of the younger face had been.
(d) The ofyf (older face with younger features) version was similar to yfof above except that the older face was the recipient of donor features from the younger face. This version was intended to present all the information available in the older face except the specified features.

In order to substitute the eyes, nose, and mouth, the donor features were pasted onto the recipient face by using the cloning tool available in Adobe Photoshop. The location of the substitution was made with precision by first selecting the same point in each of the source images and 'cloning' from that point. The inside corner of each eye was selected in turn and both eyes were copied and pasted onto the recipient face. For the nose, on a vertical plane between the eyes, the midpoint served as the starting



(a) (b) (c) (d)

Figure 1. Example of versions of a face: (a) younger (original face); (b) older (original face); (c) yfof (young face with older features); and (d) ofyf (older face with younger features).

point for cloning the 'new' nose. As with the previous features, the mouth area was cloned from the midpoint between the two lips until the 'new' mouth was completely pasted.

2.1.4 *Procedure.* Each subject saw the twenty-four images sequentially. A different random order was used for each subject. Subjects were told that the experimenter was interested in finding out how people decide how old a face is. For each face subjects were asked to write down the apparent age in years. Subjects were told that, although some faces might seem to recur in the series, none was identical to any other and they were asked to rate each face independently of any other. They were asked to take as long as they considered necessary to complete the task.

2.2 *Results*

2.2.1 *Treatment of results.* The raw data consisted of each subject's estimated ages (in years) for the twenty-four faces. Mean estimated ages were calculated for each permutation of face and face version (younger, yfof, ofyf, older). These means and their corresponding standard deviations are shown in table 1.

Table 1. Mean estimated ages and effect of features for each of the four versions of faces in experiment 1: younger, older, yfof (younger face, older features), and ofyf (older face, younger features). Standard deviations are shown in parentheses. The percentage effect of features for the yfof and ofyf versions are the mean estimated ages of the yfof and ofyf versions as percentages of the difference between the age estimates of the versions from which they are derived A plus sign indicates an overestimate, a minus sign an underestimate. The mean effect of features is the average of the absolute values of these percentages (see section 2.2.3.2).

	Face version				
	younger	yfof		ofyf	older
Male 1–5 years					
Estimated age	0.67	2.12		4.07	5.83
	(0.46)	(2.68)		(2.27)	(1.25)
Percentage effect		+28%		−34%	
Mean effect			32%		
Female 1–5 years					
Estimated age	1.49	3.40		4.19	5.43
	(1.12)	(1.35)		(1.33)	(1.57)
Percentage effect		+48%		−31%	
Mean effect			40%		
Male 4–10 years					
Estimated age	5.68	9.63		8.05	11.49
	(1.93)	(2.27)		(2.49)	(1.95)
Percentage effect		+68%		−59%	
Mean effect			64%		
Female 3–9 years					
Estimated age	4.30	7.18		9.95	10.85
	(1.57)	(2.48)		(1.32)	(1.81)
Percentage effect		+44%		−14%	
Mean effect			29%		
Male 10–40 years					
Estimated age	10.35	16.75		38.23	42.55
	(2.41)	(3.57)		(5.76)	(2.97)
Percentage effect		+20%		−13%	
Mean effect			17%		
Female 20–65 years					
Estimated age	35.58	49.23		66.50	75.13
	(8.71)	(8.20)		(9.42)	(7.58)
Percentage effect		+35%		−50%	
Mean effect			43%		

A two-way ANOVA with repeated measures was performed on these age estimates (six levels of face and four levels of version of face). This revealed a highly significant main effect of face ($F_{5,195} = 2480.62$, $p < 0.0001$) and a highly significant effect of version of face ($F_{3,195} = 767.06$, $p < 0.0001$). There was also a significant interaction between face and version ($F_{15,585} = 220.91$, $p < 0.0001$). In other words, the experimental manipulations affected the apparent age of the faces used, and these effects varied according to the individual face concerned. (This is to be expected, given the very different age ranges encompassed by these faces.)

2.2.2 *The pattern of age estimates.* For each individual face, estimates of age for the manipulated versions fell between the estimates for the younger and older versions of that face. In each case the younger version was estimated as being youngest and the older face oldest. For five of the six faces the pattern was: younger < yfof < ofyf < older. For one of the faces (male 4–10 years) the pattern was: younger < ofyf < yfof < older.

Although this pattern is quite clear, it should be noted that the standard deviations associated with the means for the manipulated versions (ofyf and yfof) are often larger than for the unmanipulated ('younger' and 'older') faces (table 1). At present we have no full explanation for this phenomenon. Possibly it might reflect a difference between subjects in their sensitivity to the feature-swapping manipulation; after all, this does introduce some discrepancy between the cues to age that are potentially available to a subject, and it is possible that to the extent that different subjects pay attention to different cues (ie internal features or other cues) then the manipulations might have had greater or lesser effects on their particular age estimates. Individual differences in use of various cues to age lies beyond the scope of the current paper. (However, experiment 2 will address this cue-mismatch issue in more detail.)

2.2.3 *The effect of changing the features*

2.2.3.1 *Statistical analysis.* To explore the effect of adding older features to a younger face and vice versa, a direct comparison of the estimates for these versions was made by using a series of matched *t*-tests. The results of these are shown in table 2. For all the faces there was a significant effect of adding younger or older features to an older or younger face. If one looks at the mean estimates of age, there is a clear trend: adding older features to a younger face made that face appear older, while adding younger features to an older face made that face appear younger.

2.2.3.2 *Mean effect of features.* While there is a significant difference between estimates of age for the original and feature-substituted versions (in the direction of the donor features),

Table 2. Results of *t*-tests comparing subjects' age estimates for original and feature-substituted versions of each of the six stimulus faces used in experiment 1. The 'younger vs yfof' column shows the results for comparisons between each 'younger' version of a face and its 'yfof' version, in which the original features were replaced with features from an older version of the same face. The 'older vs ofyf' column shows the results for comparisons between each 'older' version of a face and its 'ofyf' version, in which the original features were replaced with features from a younger face. All *t*-tests have 39 degrees of freedom and were two-tailed tests.

	Younger vs yfof	Older vs ofyf
Male 1–5 years	$t = -3.15$, $p < 0.005$	$t = 5.93$, $p < 0.0001$
Female 1–5 years	$t = -10.09$, $p < 0.0001$	$t = 7.88$, $p < 0.0001$
Male 4–10 years	$t = -10.29$, $p < 0.0001$	$t = 7.07$, $p < 0.0001$
Female 3–9 years	$t = -8.23$, $p < 0.0001$	$t = 2.54$, $p < 0.02$
Male 10–40 years	$t = -11.35$, $p < 0.0001$	$t = 4.03$, $p < 0.0005$
Female 20–65 years	$t = -9.52$, $p < 0.0001$	$t = 6.68$, $p < 0.0001$

a percentage measure of the effect of features can be calculated as follows:

$$\text{Percentage effect of younger features} = \left(\frac{\text{older} - \text{ofyf}}{\text{older} - \text{younger}}\right) \times 100\,,$$

$$\text{Percentage effect of older features} = \left(\frac{\text{yfof} - \text{younger}}{\text{older} - \text{younger}}\right) \times 100\,.$$

For example, for the 'male 4–10 years' faces, the 'age space' can be defined as the difference between age estimates for the older and younger faces (11.49 years $-$ 5.68 years $=$ 5.81 years). By taking the estimates for the older face with younger features away from those estimates for the older (unmanipulated) face (8.05 years $-$ 11.49 years $= -3.44$ years) and dividing this value by the age space ($-3.44/5.81 = -59\%$) a measure of the effect of the younger features can be obtained. Thus, in this particular case, adding younger features to an older version affected age judgments by 59%. Similar calculations were made for all faces (see table 1 and figure 2). With all faces taken together, the overall effect of feature substitution was 39%. For each face the effect varied from 17% to 64%.

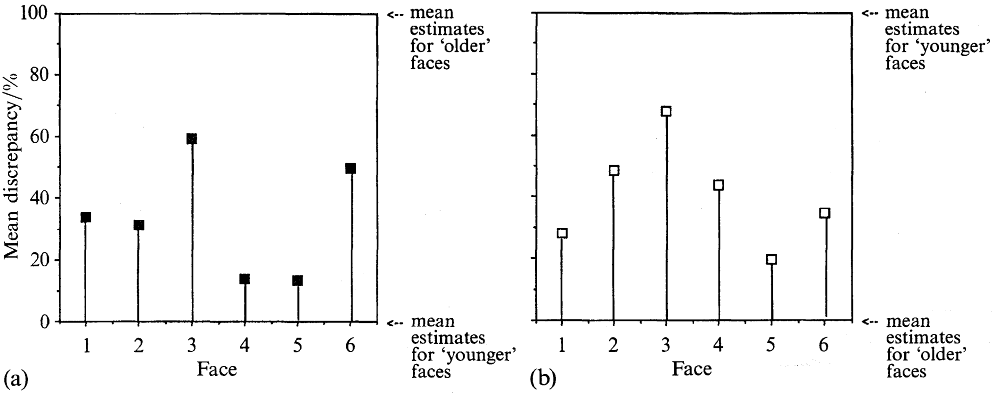


Figure 2. (a) Age estimates for each of the six stimulus faces used in experiment 1 when older features are added to a younger face. These are expressed as discrepancies, that is, percentages of the 'age space', the difference between the mean age estimates for the older (100% on the vertical axis) and younger (0% on the vertical axis) versions of a given face. The greater the deviation of age estimates from 0%, the greater the influence of the pasted features on the apparent age of the entire face, in the sense of making the face appear older. (b) Age estimates for each of the six stimulus faces when younger features are added to an older face. As in (a), these are expressed as percentages of the age space, the difference between the mean age estimates for the younger (100% on the vertical axis) and older (0% on the vertical axis) versions of a given face. As in (a), the greater the deviation of age estimates from 0%, the greater the influence of the pasted features on the apparent age of the entire face, but this time in the sense of making the face appear younger.

2.3 *Discussion*

The aim of this experiment was to explore the influence of specific facial features (the eyes, nose, and mouth) on the processing of a face's age. The features of younger and older versions of six individual faces were substituted for each other in the same location as the recipient face. The results suggest a clear trend: adding older features to a younger face and younger features to an older face produced significantly different estimates of age compared with those given for the original faces. The difference was in the direction of the 'donor' features.

For all faces except one, these results are consistent with the hypothesis that specific feature information can be influential in conveying information about age.

However, it is clear from these results that feature information was not the *sole* determinant of apparent age. If this were so, then yfof faces would have been judged as older than ofyf faces—which was not the case for five of the six sets of faces used. (We are indebted to one of our reviewers for pointing this out.) In fact, in all but one case, the ages of the yfof faces were estimated to be *younger* than the ofyf faces, suggesting that other factors (ie cues available from the rest of the face) still exerted a considerable influence on apparent age.

One way to conceptualise the influence of substituted features on apparent age is to take the difference between the age estimates for the older and younger versions (the age space) and use the amount of any shift in age-estimation performance along this age space for the feature-substituted faces as a measure of the influence of those features. When all the stimulus faces are taken together, feature substitution produced age estimates that represented an overall shift of 39% from those estimates given for the original faces.

If the diverse age range of faces used is considered, the significant effects of feature substitution for all faces are rather surprising. If one assumes that age processing of faces requires making qualitatively and quantitatively different discriminations according to the age of the face itself, some cues may be more relevant to growth-related changes than for those faces that have stopped growing. While in this experiment we have found that feature substitution influenced the apparent age of all the faces used, it could be argued that the features themselves are intrinsically linked with other cues. For example, the features of older faces will necessarily carry some 'older' surface-texture information (eg the skin covering the nose) and this may confound measurement of the influence of the features per se. In addition, some features may be more or less important, according to the age of the face itself. For example, it is plausible that if teeth size and shape are a cue used for the age processing of faces of young children, then the mouth area may be more important for this age group.

Differential usage of cues and the possible confounding of features with other sources of information (eg skin-texture cues) causes some problems for any interpretation of the importance of the features with the faces used in this experiment. For example, consider the age estimates for the 3 – 9-year-old girl and the 4 – 10-year-old boy. As a reviewer pointed out, one might reasonably expect the age estimates for these children to be similar, given that they are both at approximately the same developmental stage. The observed differences in age estimates for these two individuals might be due to a number of factors which it was difficult to control for in this particular experiment. For example, the boy had a similar hairstyle at ages 4 and 10 years, whereas the girl's hairstyle was more appropriate to her age in both of the original stimuli (ie at ages 3 and 9 years). Consequently, it might be that the effects of the feature manipulations occurred within a more plausible context in the case of the boy than in the case of the girl.

Clearly the sample of faces was rather small to allow firm generalisations to be made regarding the influence of feature substitution in conveying information about age. The fact that the interaction between face and version was statistically significant reinforces this concern: there is the possibility that some of these effects might be due to the particular faces used. Experiment 2 was therefore performed in an attempt to address these problems and increase the generalisability of these findings by using a larger number of stimulus faces.

3 Experiment 2

In experiment 2 we attempted to extend and clarify the results of experiment 1, controlling for the potential problems already discussed by using a larger number of faces within a well-defined age range: after growth is complete and before major changes in surface cues have occurred. [Enlow (1982) suggests that major changes in skin texture and underlying musculature are generally not noticeable until about the age of 40 years.] Because of the difficulty in obtaining enough suitable stimulus faces and the problems of subjects' possible sensitivity to age-irrelevant stimulus attributes (differences between the younger and older images in terms of photographic quality, slight changes in head orientation, etc) it was decided not to use the same individual at different ages as source material but to use two different individual faces (one younger and one older) from two age ranges (18 – 22 years and 28 – 38 years) where the quality and pose of the photographs could be controlled.

Ten face pairs were used. Faces were paired according to the following criteria: they were approximately 10 years apart in age, and similar in terms of head orientation, lighting of the face, hairstyle, and facial expression. For each face pair, a morphed version was produced that represented a 50% blend of both faces. This provided a standard face that served as a recipient for younger and older features. Using this method provided us with stimuli which remained naturalistic in appearance while minimising any differences in skin texture between the younger and older faces.

For each of the ten sets of faces five versions were produced. Two were unmanipulated versions ('younger' and 'older'—original younger and older faces). Three were variants of these: a blend of the younger and older faces ('morph'), morph face with younger features (mfyf), and morph face with older features (mfof). The subjects' task was to estimate the age of each of the resultant fifty faces.

Our predictions were as follows. First, as the 'morph' version is a blend of the younger and older source faces from which it is derived, then it was expected that the estimates of age for this version would fall between those estimates given for the younger and older original faces. Second, as suggested by experiment 1, if feature information is influential in the processing of a face's age, then the morph face with young features (mfyf) and the morph face with older features (mfof) would be rated as being younger and older respectively than the morph face. This being the case, the pattern of age-estimation performance for all faces should be: younger < mfyf < morph < mfof < older.

3.1 *Method*

3.1.1 *Subjects.* Forty subjects aged between 18 and 40 years took part. Half of the subjects were male, half were female. All subjects were unaware of the purpose of the experiment.

3.1.2 *Design.* A within-subjects design was used, with all subjects participating in all five conditions of the experiment: (a) younger, (b) morph face with younger features, (c) morph face, (d) morph face with older features, and (e) older.

There were two independent variables. The first was the face pair. Ten face pairs were used, each comprising one younger original and one older original face. The second independent variable was the version of target face. Five different versions of each set of face pairs were used, two derived from photographs and three manipulated versions (see below for definitions). The dependent variable was the subject's estimate of the age of each face, in years.

3.1.3 *Apparatus and materials.* Twenty photographs of males were taken (ten photographs for each of two age categories, younger and older). These were scanned into Adobe Photoshop and Gryphon morphing image-editing software on a Power Macintosh 7100/66 by using a Macintosh OneScanner set for 256 grey levels and 150 dots inch^{-1}. The twenty photographs were made into ten face pairs by matching each younger face with an older one. For each of the ten sets five versions were produced. Each set consisted of the original younger and original older face plus three variants: a morphed face (morph), a morphed face with older features (mfof), and a morphed face with younger features (mfyf). After manipulation the images were output as fifty A4-sized prints on a BLP Eclipse 8 laser printer. An example of a set of five versions can be seen in figure 3. The different versions were as follows.

(a) The 'younger' version (between ages 18 and 22 years). This version was an unaltered print from the original scanned photograph.

(b) The 'older' version (between ages 28 and 38 years) was an unaltered print from the original scanned photograph.

(c) The 'morph' version was intended to provide a blend of the younger and older faces that was intermediate between them in terms of head orientation, lighting, hair-style, etc. The sources of age information available in the two original faces were also

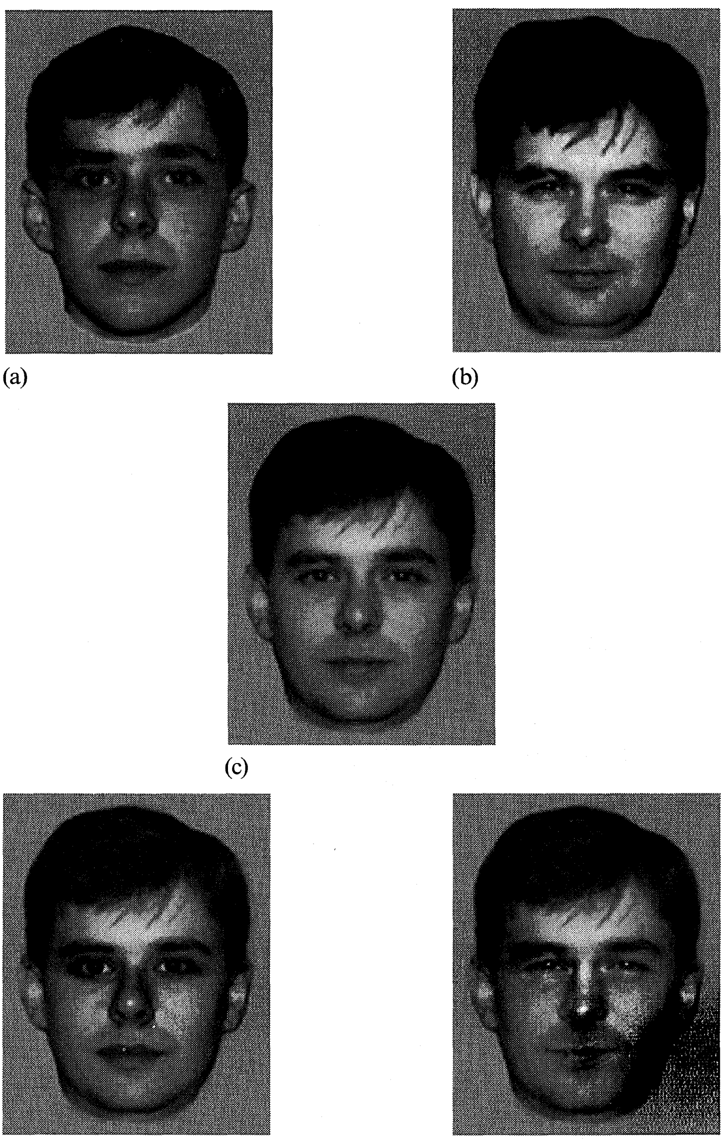


Figure 3. Example of versions of a face: (a) younger; (b) older; (c) morph; (d) morph with younger features; and (e) morph with older features.

blended by the morphing process, thus producing a face with age information that was intermediate between the 'younger' and 'older' versions.
(d) The mfof version was intended to present all the information available in the morph face except for specific features. The eyes, nose, and mouth from the younger version were pasted onto the morph face. The donor features were placed in the same location as the corresponding features in the recipient image.
(e) The mfyf version was similar to mfof above except that the donor features were from the older face. This version was intended to present all the information available in the morph face and the version mfof except for the specified features.

3.1.4 *Procedure.* The procedure was identical to that of experiment 1, except that subjects saw fifty images.

3.2 *Results*

3.2.1 *Treatment of results.* The raw data consisted of each subject's estimated ages (in years) for the fifty faces. Mean estimated ages were calculated for each permutation of face and face version (younger, mfyf, morph, mfof, older). These means and their corresponding standard deviations are shown in table 3.

A two-way ANOVA with repeated measures was performed on the data (with ten levels of face pair and five levels of version of face). This revealed a significant main effect of face pair ($F_{9,390} = 30.64$, $p < 0.0001$) and a highly significant effect of version of face ($F_{4,1560} = 1490.77$, $p < 0.0001$). There was also a significant interaction between face pair and version ($F_{36,1560} = 9.48$, $p < 0.0001$).

Table 3. Mean estimated ages for the five versions of each of the ten face pairs used in experiment 2 (younger, mfyf, morph, mfof, and older). Standard deviations are in parentheses. For each face pair, the 'morph' was a 50% blend between the 'younger' and 'older' faces in that pair, and the mfyf and mfof versions consisted of this morph with features replaced with those from the 'younger' or 'older' face, respectively. See sections 2.2.3.2 and 3.2.3.2 for details of how the mean effect of features was calculated.

Face pair	Face version					Mean effect of features
	younger	mfyf	morph	mfof	older	
1	20.48	22.35	26.03	31.68	35.73	62%
	(4.72)	(4.74)	(3.68)	(4.03)	(5.67)	
2	23.28	28.35	29.38	32.25	37.85	26%
	(4.47)	(5.31)	(4.53)	(5.14)	(5.05)	
3	18.28	20.45	24.58	27.95	32.80	53%
	(2.42)	(4.48)	(3.85)	(4.48)	(4.86)	
4	18.93	22.20	25.30	28.05	36.20	39%
	(2.75)	(2.96)	(3.60)	(5.31)	(4.87)	
5	18.98	23.15	27.35	33.68	39.10	52%
	(2.94)	(4.23)	(3.06)	(4.76)	(4.83)	
6	22.95	29.83	32.40	37.05	41.38	39%
	(5.50)	(4.71)	(5.32)	(5.78)	(4.37)	
7	18.23	23.83	26.58	29.98	40.45	35%
	(2.87)	(3.92)	(3.72)	(5.52)	(5.57)	
8	17.45	21.70	25.20	31.05	37.30	47%
	(2.25)	(4.15)	(5.56)	(5.01)	(6.05)	
9	20.93	27.93	34.90	37.15	44.08	38%
	(3.72)	(4.39)	(6.30)	(6.11)	(5.76)	
10	17.08	28.00	32.20	37.48	39.75	49%
	(3.25)	(5.80)	(4.71)	(5.61)	(4.86)	

3.2.2 *The pattern of age estimates.* For each face pair the mean estimates of age for the manipulated versions fell between the estimates for the 'younger' and 'older' versions. In each case the 'younger' version was estimated as being youngest and the 'older' face oldest. In addition, the estimated ages for the blended version 'morph' were rated as being in the middle. Clearly, there are differences between the face pairs (as reflected in the highly significant interaction between face pair and version). We have no interesting explanation for these; presumably they reflect differences between the faces used, in terms of their susceptibility to the manipulations used. However, the same trend was found for all ten face pairs: younger < mfyf < morph < mfof < older.

Figure 4 shows the mean estimates for the manipulated (mfyf, morph, and mfof) as percentages of age space between estimates for 'younger' and 'older' versions. A mean value of 0% would reflect estimates of age equivalent to those for the 'younger' version. In contrast, a mean value of 100% would reflect equivalence with the estimates for the 'older' version. In general terms, the greater the percentage between mfyf and mfof, the greater the effect of feature information on subjects' age estimates.

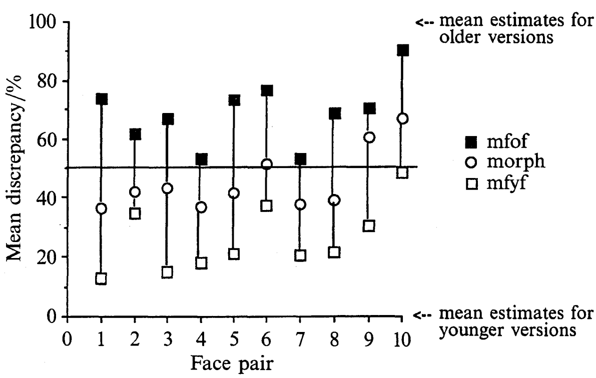


Figure 4. Age estimates for each of the ten stimulus face pairs used in experiment 2. As in figure 2, these are expressed as discrepancies, that is, percentages of the age space (the difference between the mean age estimates for younger and older stimuli in the face pairs presented). The graph shows mean age estimates for mfof (morphed face with older features), morph, and mfyf (morphed face with younger features) stimuli. If age estimates were determined by a 'blending' of all age information contained in the younger and older faces of a pair, then age estimates would fall on the 50% line displayed. As can be seen from the graph, most morphs were perceived as being younger than one would expect from such a simple blending process; the mfof faces were all perceived as being older than the morphs (reflecting the influence of adding older features); and the mfyf faces were all seen as being younger than the morphs (reflecting the influence of adding younger features).

3.2.3 *The effect of changing the features*

3.2.3.1 *Statistical analysis.* To explore the effect of adding older and younger features to a 'morph' face, a direct comparison of the age estimates for the versions 'mfyf' and 'mfof' with those for the morph face was made, by using a series of matched *t*-tests. For nineteen out of twenty comparisons, there was a significant effect of adding younger or older features to the morph face: age estimates were significantly different from those given to the morphed face, and were biased towards the age of the face from which the features came. These results are shown in table 4.

3.2.3.2 *Mean effect of features.* By using a similar method to that in experiment 1, by taking the difference between the morph faces (recipient face for both the donor features) and the 'younger' and 'older' versions, the age space within which the manipulated versions can be placed was calculated. Any effect of experimental manipulation of over 50% would reflect the feature information as being more influential than any other source(s) of information for the task at hand. The absence of any effect would

Table 4. Results of t-tests comparing subjects' age estimates for morphed and feature-substituted versions of each of the ten stimulus face pairs used in experiment 2. The 'morph vs mfyf' column shows the results for comparisons between age estimates given to each face-pair morph and the same morph with younger features pasted into it. The 'morph vs mfof' column shows the results for comparisons between age estimates given to each face-pair morph and the same morph with older features pasted into it. All t-tests have 39 degrees of freedom and were two-tailed tests.

Face pair	Morph vs mfyf	Morph vs mfof
1	$t = -5.85$, $p < 0.0001$	$t = 9.83$, $p < 0.0001$
2	$t = -1.57$, ns	$t = 3.93$, $p < 0.0005$
3	$t = -6.59$, $p < 0.0001$	$t = 5.48$, $p < 0.0001$
4	$t = -5.50$, $p < 0.0001$	$t = 4.43$, $p < 0.0002$
5	$t = -7.80$, $p < 0.0001$	$t = 8.65$, $p < 0.0001$
6	$t = -6.30$, $p < 0.0001$	$t = 6.53$, $p < 0.0001$
7	$t = -4.52$, $p < 0.0002$	$t = 5.00$, $p < 0.0001$
8	$t = -6.00$, $p < 0.0001$	$t = 6.92$, $p < 0.0001$
9	$t = -7.84$, $p < 0.0001$	$t = 2.69$, $p < 0.01$
10	$t = -5.11$, $p < 0.0001$	$t = 6.41$, $p < 0.0001$

indicate that the features had no influence on age estimates. If all face pairs are taken together, the mean effect of substituting features was a 44% shift in age estimates from the original source faces. By comparing performance between the face pairs, the mean overall effect ranged from 26% to the highest effect which was 62% (see table 4).

3.3 *Discussion*

The subjects' task in experiment 2 was to estimate the age of ten sets of faces. For each set there were two original faces (one younger and one older) and three manipulated versions, morph, mfyf, and mfof. The blended (younger with older) morph face was used as a basis for both sets of features for substitution, and as a measuring point from which the effect of the feature-substituted versions could be ascertained. For all face sets, the estimated ages for the morph versions were rated as being between the estimates for the younger and older source faces. In addition, the estimates of age for the feature-substituted versions were in the predicted direction relative to the morph face. These results confirm and extend those found in experiment 1: when using a larger number of faces within a defined age range, substituting features of a different age produced a significant influence on age estimates. For this set of faces, apparent age is strongly influenced by the age of the substituted features.

As with experiment 1, by taking the age estimates for the younger and older (original) faces and the morph face to represent the maximum amount of age space that is available to the perceiver, a measure of the effect of features on apparent age was made by calculating the amount of shift in estimates for the morph faces and the morph faces with substituted features. In experiment 1, the mean effect was 39%; the lowest effect was 17% and the highest was 64%. The results found in experiment 2 are broadly similar: the overall effect of substituting features produced a mean shift in age space of 44%, with a lowest effect of 26%, and a highest effect of 62%. For one of the faces in the first experiment and three out of ten sets of faces used in the second experiment, the feature information was more influential than any other source(s) of available information.

4 General discussion

In previous studies the importance of specific cues in the age processing of faces has been emphasised: cardioidal strain (Pittenger and Shaw 1975; Pittenger et al 1979; Mark and Todd 1983); the spatial interrelationship of the features (George and Hole 1995; Bruce et al 1991); and skin information (in the form of wrinkles; Mark et al 1980).

The present study provides an indirect measure of these cues for age perception. The results found here suggest that although these cues, either individually or in combination, are important cues to age, they are certainly not the only type of information that is used. On the contrary, our results suggest that the features themselves may be an influential source of information about age. When subjects are given a choice of cue(s) to use, by placing feature information in opposition to other cues, they are strongly influenced by the age of the features present. For example, in the feature-substituted versions, the new features were in the same location as the old ones: if cardioidal-strain-level information were the paramount cue to age, then subjects' age estimates should have been based on the feature position, not the features themselves. Equally, if skin information (from facial regions other than the features themselves) were the most salient cue for determining a face's age, then feature substitution should have had little effect on age estimates.

One conclusion is that studies which show that subjects are sensitive to cardioidal-strain level and skin cues (at least, nonfeatural skin cues), have led to an overemphasis on the importance that these types of information play in age perception. [It is only fair to point out that Pittenger and his colleagues have always been circumspect in their interpretation of cardioidal strain as a cue to age; in fact Mark et al (1980) stress the complexity of the interaction between various cues to age. However, their demonstrations of the effects of cardioidal strain are so compelling that it has been easy to overlook the role of other sources of information about age. Possibly an impression that the question of how age is perceived has been 'settled' is the reason why so little research has been undertaken on this topic, in stark contrast to the huge amount of research that has been performed on face recognition in recent years.]

The results from the present study support and extend our previous findings (George and Hole 1995) by suggesting that the internal features are an influential source of age information for faces of all ages. While global shape transformation (such as the cardioidal-strain transformation researched by Pittenger, Shaw, and their colleagues) is certainly an important source of information during growth, it certainly is not the only source of age information available to the perceiver, either for faces that are growing or for faces that have finished growing. While we showed (George and Hole 1995) that age perception could be estimated accurately solely on the basis of the facial region containing the internal features, we did not provide information on *which* of the cues present within this region were used, since many were retained in our 'internal-features' stimuli. Subjects could have been basing their judgments on the details of the features themselves; on the texture of skin regions (parts of the cheeks) that were visible around the features; or on the spatial arrangements of the features (either in a direct sense or as an indirect means by which to estimate the degree of cardioidal strain of the entire head).

The present experiments constitute an advance because they enable these candidate sources of perceptual information to be decided between, by separating the features from their immediate surroundings and from their original configuration. When features from a face of one age are placed into the locations of the original features of a face of a different age, subjects are influenced by this manipulation—not solely by more-global properties of the recipient face. This is not what one would predict from an account of age perception based primarily on cardioidal strain or other *global* transformations, and it is not what one would expect from an account of age perception based only on *global* textural characteristics such as wrinkling. In short, our findings to date suggest that *local* sources of information (eg featural details) may be used as a basis for accurate estimates of age.

However, this 'local' versus 'global' distinction may not be as clear-cut a distinction as it at first sight appears, and we certainly are not proposing any simplistic dichotomy between the two types of cue. We cannot exclude the possibility that our apparently 'local' alterations to the face (by feature substitution) produced global changes—not just in the sense of altering the 'configuration' between the features themselves, but in the sense of altering the three-dimensional interpretation of the entire craniofacial complex. Our *finding* is that feature substitution alters the apparent age of a face; the *interpretation* of this is considerably more complex than it appears at first sight. Paradoxically, it may be that feature substitution has its effects because it changes the viewer's interpretation of the three-dimensional shape of the head, and thus is actually a demonstration of the importance of *global* cues to age!

Whatever one's preferred interpretation, a local/global distinction is less important than an appreciation that *all* age changes—whether local or global—reflect the underlying biological changes produced by growth and ageing. Apparently 'local' changes in size and shape of features may reflect changes in the underlying global morphology, just as much as more obviously 'global' changes due to cardioidal strain—one must not lose sight of Bruce's (1988) insight that "faces grow on heads" (page 118) and that featural changes may reflect this fact.

In the context of the present study, precisely what it is about the eyes, nose, and mouth that conveys information about age is still not directly addressed, and therefore can only be speculated upon. It remains to be determined whether all of these features are important or whether some are more influential than others. For example, would age perception be accurate on the basis of the eyes alone? Second, it also remains to elucidate precisely which aspect of the feature(s) concerned is important for age perception: size, shape, or texture are obvious candidates.

However, by taking our other findings (George and Hole 1995) together with findings from a study which we have recently completed (George and Hole, submitted) we can provide a tentative answer to the second question. By use of thresholding in the former study and low-pass spatial-frequency filtering (blurring) in the latter, we have found that when detailed feature information is reduced, the accuracy of age perception is impaired. Although the internal facial features are sufficient for accurate age perception, it seems that information about the surface detail of these features has to accompany information about their shape for this to be accomplished.

Finally, previous researchers have made a distinction between the types of information used in the processing of familiar and unfamiliar faces. Ellis et al (1979) showed that while internal and external features of faces were used as cues for recognising previously presented but otherwise unfamiliar faces, recognition of familiar faces was relatively more successful from internal rather than external features. Young et al (1985) asked subjects to compare different views of faces, and found that matching internal features of familiar faces was significantly easier than matching internal features of unfamiliar faces. This suggests that when the task is one of recognition, the internal features may be more important in the processing of familiar faces. In contrast to this, the present study supports the hypothesis that for age processing of unfamiliar faces, the internal features are important. This is further, albeit indirect, evidence that information from faces is used for a variety of tasks (individual recognition, expression analysis, lipreading, and age and gender perception) and that the processes underlying these different tasks may have different requirements and are largely dissociable from each other (reviewed in Bruce 1988; see also Humphreys et al 1993).

In conclusion, in this study we initially set out to explore the influence of the features in assessing a face's age. In the first experiment it was clear that the features were influential for faces varying in age from 1 to 65 years. Using more faces within a well-defined age range (after growth is complete and before major skin changes have occurred), in experiment 2 we confirmed and extended these results. More obviously global sources of information (as highlighted by previous research, eg cardioidal-strain level, surface texture, or the spatial interrelationship of the features) are almost certainly implicated in the age processing of faces. However, the effects of feature substitution on apparent age raise the possibility that local information about age can be obtained from the features themselves and may have a considerable influence on age estimates—not only for faces that have finished growing but for those faces that are in the process of growth.

Acknowledgements. We would like to thank the children and adults who gave permission for us to use their faces in these experiments, and our anonymous reviewers for their helpful and thoughtful comments and suggestions on earlier drafts of this paper.

References

Bartlett J C, Searcy J, 1993 "Inversion and configuration of faces" *Cognitive Psychology* **25** 281 – 316

Bruce V, 1988 *Recognising Faces* (London: Lawrence Erlbaum Associates)

Bruce V, Burton M, Doyle T, Dench N, 1989 "Further experiments on the perception of growth in three dimensions" *Perception & Psychophysics* **46** 528 – 536

Bruce V, Doyle T, Dench N, Burton M, 1991 "Remembering facial configurations" *Cognition* **38** 109 – 144

Carey S, Diamond R, 1977 "From piecemeal to configurational representation of faces" *Science* **195** 312 – 314

Davies G M, Ellis H D, Shepherd J W (Eds), 1981 *Perceiving and Remembering Faces* (London: Academic Press)

Ellis H D, Shepherd J W, Davies G M, 1979 "Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition" *Perception* **8** 431 – 439

Enlow D, 1968 *The Human Face: An Account of the Postnatal Growth and Development of the Craniofacial Skeleton* (London: Harper and Row)

Enlow D, 1982 *The Handbook of Facial Growth* second edition (Philadelphia, PA: W B Saunders Company)

Fraser K I, Parker D, 1986 "Reaction time measures of feature saliency in a perceptual integration task", in *Aspects of Face Processing* Eds H D Ellis, M A Jeeves, F Newcombe, A Young (Dordrecht: Martinus Nijhoff) pp 45 – 52

George P A, Hole G J, 1995 "Factors influencing the accuracy of age estimates of unfamiliar faces" *Perception* **24** 1059 – 1073

George P A, Hole G J (submitted), "The role of spatial and surface cues in the age-processing of unfamiliar faces"

Humphreys G W, Donnelly N, Riddoch M J, 1993 "Expression is computed separately from facial identity and it is computed separately for moving and static faces: neuropsychological evidence" *Neuropsychologia* **31** 173 – 181

Jones G, Smith P K, 1984 "The eyes have it: young children's discrimination of age in masked and unmasked facial photographs" *Journal of Experimental Child Psychology* **38** 328 – 337

Liggett J, 1974 *The Human Face* (London: Constable)

Luria S M, Strauss M S, 1978 "Comparison of eye movements over faces in photographic positives and negatives" *Perception* **7** 349 – 358

Mark L S, Pittenger J B, Hines H, Carello C, Shaw R E, Todd J T, 1980 "Wrinkling and head shape as co-ordinated sources of age-level information" *Perception & Psychophysics* **27** 117 – 124

Mark L S, Todd J T, 1983 "The perception of growth in three dimensions" *Perception & Psychophysics* **33** 193 – 196

Mark L S, Todd J T, 1985 "Describing perceptual information about human growth in terms of geometric invariants" *Perception & Psychophysics* **37** 245 – 256

Perrett D I, Mistlin A J, Chitty A J, 1987 "Visual neurones responsive to faces" *Trends in Neurosciences* **10** 358 – 364

Pittenger J B, Shaw R E, 1975 "Ageing faces as visceral-elastic events: Implications for a theory of nonrigid shape perception" *Journal of Experimental Psychology: Human Perception and Performance* **1** 374 – 382

Pittenger J B, Shaw R E, Mark L S, 1979 "Perceptual information for the age-level of faces as a higher-order invariant of growth" *Journal of Experimental Psychology: Human Perception and Performance* **5** 478 – 493

Rhodes G, 1995 "Face recognition and configural coding", in *Cognitive and Computational Aspects of Face Recognition* Ed. T Valentine (Andover, Hants: Routledge, Chapman and Hall) pp 47 – 68

Rhodes G, Brake S, Atkinson A P, 1993 "What's lost in inverted faces?" *Cognition* **47** 25 – 57

Shepherd J W, Davies G M, Ellis H D, 1981 "Studies of cue saliency", in *Perceiving and Remembering Faces* Eds G Davies, H Ellis, J Shepherd (London: Academic Press) pp 105 – 131
Smith D W, 1978 *Growth in the Biological Ages of Man: From Conception through to Old Age* second edition, Eds D W Smith, E L Bierman, N M Robinson (Philadelphia, PA: W B Saunders)
Young A W, Hay D C, McWeeny K H, Flude B M, Ellis A W, 1985 "Matching familiar and unfamiliar faces on internal and external features" *Perception* **14** 737 – 746
Viidik A, 1973 "Functional properties of collagenous tissues", in *International Review of Connective Tissue Research* volume 6, Eds D A Hall, D S Jackson (London: Academic Press)